

Title:

State-Wide Performance Criteria for International Safeguards

Author(s):

Kory W. Budlong Sylvester, Joseph F. Pilat, and William D. Stanbro

Submitted to:

<http://lib-www.lanl.gov/la-pubs/00796148.pdf>

State-Wide Performance Criteria for International Safeguards

Kory W. Budlong Sylvester, Joseph F. Pilat, and William D. Stanbro

Los Alamos National Laboratory

Los Alamos, NM 87545 USA

(505) 665-9200 FAX: (505) 667-7626

Abstract

Traditionally, the International Atomic Energy Agency (IAEA) has relied upon prescriptive criteria to guide safeguards implementation. The prospect of replacing prescriptive safeguards criteria with more flexible performance criteria would constitute a structural change in safeguards and raises several important questions. Performance criteria imply that while safeguards goals will be fixed, the means of attaining those goals will not be explicitly prescribed. What would the performance objectives be under such a system? How would they be formulated? How would performance be linked to higher level safeguards objectives? How would safeguards performance be measured State-wide?

The implementation of safeguards under performance criteria would also signal a dramatic change in the manner the Agency does business. A higher degree of flexibility could, in principle, produce greater effectiveness and efficiency, but would come with a need for increased Agency responsibility in practice. To the extent that reliance on prescriptive criteria decreases, the burden of justifying actions and ensuring their transparency will rise. Would there need to be limits to safeguards implementation? What would be the basis for setting such limits? This paper addresses these and other issues and questions relating to both the formulation and the implementation of performance-based criteria.

Introduction

The possibility of utilizing performance criteria in the context of integrated safeguards has been raised previously.¹ In this paper the general components of performance criteria are revisited and key elements are expanded upon.

What Are Performance Criteria?

Performance criteria offer an alternative to prescriptive means of meeting organizational goals, in this case of IAEA safeguards. Rather than specifying in detail the exact set of activities that must be performed for goal attainment, multiple means for meeting the goals are allowed. The emphasis is placed on meeting functional requirements rather than the means by which they are attained, on output rather than input.

In the process, organizational goals are translated into technical parameters against which the functional effectiveness of a set of actions can be measured. In a regulatory context, an example would be setting safety limits (e.g., expected core damage frequency) and allowing the regulated party flexibility in terms of how those limits are met. As long as it can be demonstrated that the limits are being met reliably, the approach used can be deemed acceptable.

Elements of a performance-based approach can be found in the design approach underlying current safeguards. In current safeguards, IAEA detection goals play an important role. Technical parameters such as significant quantities, detection time and detection probability have been

established and guide current safeguards implementation. The difference is, in a performance-based approach, the process stops short of prescribing actions.

Performance criteria would be wholly consistent with the IAEA's non-discriminatory principle governing the development of integrated safeguards. In an August 2000 General Conference paper the Agency noted "... although the measures actually used in individual States may differ, the same technical objectives must be pursued in all States with comparable safeguards obligations..."² This requirement could be supported through the development and use of performance criteria.

A major driver for moving away from prescriptive-based strategies and to more performance-based approaches is cost effectiveness. It is easy to see that a one-size-fits-all approach can, in some cases, result in unintended outcomes. For example, if prescriptive criteria are followed too strictly, effort could be expended in areas of little importance to goal attainment. Furthermore, if case-specific factors are not properly taken into account, areas of significance can be missed. Performance criteria, when used properly, can better support the efficient allocation of limited resources.

International Safeguards Objectives

Technical performance criteria should always be directly traceable to high-level objectives. For international safeguards, performance criteria should reflect the high-level objectives that have directed the IAEA since its inception: verify compliance with safeguards agreements and provide for timely detection of violations.

Safeguards activities must allow the IAEA to draw conclusions regarding the completeness and correctness of States' safeguards declarations. Full-scope safeguards require States to make complete declarations regarding their nuclear-related activities and to employ declared material and facilities only in peaceful uses. A major function of safeguards is to provide assurances that these commitments are adhered to.

Another functional objective of safeguards is to provide timely warning. Should a State decide to violate its safeguards agreements, safeguards are intended to provide timely detection. In providing this risk of early detection, the intent is to deter proliferation.

Taken together, verification and timely warning are two functional objectives of safeguards that can provide the basis for performance criteria development.

Performance Measures

In developing performance criteria, we must first translate functional objectives into performance measures. These performance measures must capture the intent of safeguards and provide a means for measuring the degree to which objectives are being met. A number of different performance measures may be necessary to capture all relevant factors. Performance criteria are then created by specifying the values required across the set of performance measures deemed necessary for goal attainment.

As an example, the capability to detect actions that would constitute violation of safeguards agreements is a common thread between the objectives of verification and timely warning. Verification is achieved by testing the hypothesis that safeguards violations have occurred. "If

careful verification activities lead to the conclusion that the diversion hypothesis cannot be substantiated, then it can be concluded with a high level of confidence that in fact no diversion has occurred.”³ Thorough verification activities are designed to detect noncompliance, test proliferation hypotheses and thereby provide assurances.

In a similar manner, the functional objective of timely warning is achieved via detection capabilities. Safeguards have historically been designed to detect the diversion of significant quantities of nuclear material roughly within conversion time. Therefore detection capability would be an important performance measure for safeguards.

Considering the above discussion, a performance measure involving detection capability could be developed in the following manner. The IAEA seeks to detect violations of safeguards agreements to both provide confidence and timely warning. There are a number of actions that would constitute a violation, e.g, diversion of declared spent fuel, construction of an undeclared hot cell, etc. The proficiency of a safeguards system in detecting each action within a given amount of time is a meaningful measure of its performance.

Performance Criteria

Safeguards performance criteria could be expressed at a number of levels of analysis and with varying degrees of generality. Using the performance measure of detection capability, several example approaches are described below.

Example 1

Using detection capability as a performance measure, criteria can be set to ensure timeliness and verification objectives are being met. As a hypothetical performance criterion, the safeguards system could be required to:

provide a medium probability of detecting the diversion of 1 SQ of declared plutonium oxide within one month for timeliness purposes, and provide a high probability of detecting the diversion of 1 SQ of declared plutonium oxide within one year (i.e., on an annual basis) for verification purposes.

If the safeguards system can be demonstrated to reliably provide such detection capabilities, it can be said to satisfy these performance criteria. Again, flexibility is achieved (albeit to a limited degree in this case) by not prescribing how the criteria should be met, but requiring only that the standard for performance be met.

In the above example, the criteria were specified for a specific material, plutonium oxide. No matter where the material is located in a State, a proposed safeguards approach would be designed with these performance criteria in mind. The same criterion could be further generalized to apply to any direct-use material. (Such generalizations are seen in the timeliness goals of the current criteria.)

It is notable that in this formulation the same technical parameters underlying current safeguards appear. Significant quantities, detection time and detection probability all play a role. In the performance criterion, the interdependence of these factors is made explicit. It is often the case presently that timeliness goals are referred to separately from detection goals. In this formulation,

each element must be included to specify fully and give meaning to the criterion. As mentioned, this formulation does offer limited additional flexibility in comparison to traditional prescriptive safeguards criteria.

Example 2

Criteria can be expressed at a higher level of analysis as well. Rather than specifying criteria for individual proliferator actions, criteria could be developed for a set of actions leading to weapons-usable material. For example, the safeguards system could be required to:

provide a medium probability of detection prior to the acquisition of 1 SQ of separated Pu or HEU for timeliness purposes, and provide a high probability of detection within a year after this time for verification purposes.⁴

In this expression of performance criteria, no individual actions are identified. The criteria apply uniformly across all acquisition paths. (An acquisition path (AP) is a set of proliferation actions that lead to the acquisition of weapons-usable material.)

This example, in practice, may not be deemed appropriate. Requiring a safeguards system to provide the same level of detection across all APs may not be feasible or desirable given resource limitations and other factors.

Example 3

Setting separate criteria for a few classes of paths may be more fitting. For example, it might be reasonable to focus safeguards resources on APs that were cheaper and technically easier to implement. In this case criteria could take the following form:

for acquisition paths where only declared material and facilities are utilized, provide a medium probability of detection prior to the acquisition of 1 SQ of separated Pu or HEU for timeliness purposes, and provide a high probability of detection within a year after this time for verification purposes.

In this approach, performance criteria for a few ‘classes’ of APs could be developed. They would, in principle, span all possible APs when taken together; such an approach allows one to avoid the problem of treating all paths equally.

The resulting criteria for each class of AP could be listed in a simple summary table. Each class would have a required detection probability within a defined time period for timeliness and for verification purposes. An illustrative example is shown in Table 1.

If so desired, performance criteria established at the acquisition path level could be translated into lower level criteria. That is to say, performance criteria of the type described in Example 1 above could be extracted from path-level criteria. This might be useful in practice, for example, if safeguards at declared facilities of the same type are to be standardized across all States.

Table 1 Illustrative Acquisition Path Performance Criteria

Class of AP	Detection Criteria for Timeliness Objective*	Detection Criteria for Verification Purposes**
Acquisition Paths of Type 1	High	High
Acquisition Paths of Type 2	Medium	High
Acquisition Paths of Type 3	Medium	Medium
etc.

*Detection occurs prior to the end of an acquisition path.

**Detection occurs within one year after the end of the acquisition path.

Defining criteria at the path level rather than at the level of individual actions potentially simplifies the problem. It may not be necessary to define detection criteria for all possible proliferation actions. We need only extract such criteria when they are needed for some practical purpose such as the one mentioned above.

No matter how the criteria are expressed, they should be internally consistent. That is to say, the detection criteria for individual proliferation actions, entire acquisition paths and any State-wide criteria should not conflict.

Establishing criteria at the path level allows for significant flexibility. What matters is the strength of coverage of the path, not coverage of an individual step on the path. Therefore not only can the safeguards used to detect individual proliferation actions be varied, but emphasis can be placed on different portions of an AP as long as the same path level performance criteria are satisfied.

Setting Performance Criteria

No matter the form of the criteria, the required level of performance is obviously a policy decision. Only the Agency can set such requirements. In doing so they are in effect defining what is meant by timely detection and the level of assurance associated with verification activities. Performance criteria must be usable by the Agency. In thinking about how to formulate criteria, it is useful to consider the possible role they would play in Agency decisionmaking.

The IAEA has articulated several basic objectives of integrated safeguards.⁵ The Agency has said that integrated safeguards approaches should achieve “the maximum effectiveness and efficiency within available resources...” In doing so, the approaches must, amongst other things, be non-discriminatory, take into account available information for the State as a whole, and provide coverage of all plausible APs.

In reviewing Agency integrated safeguards objectives, setting performance criteria at the path level appears to offer several potential practical advantages. Acquisition paths play an explicit or implicit role in several of the Agency’s objectives. Effectiveness in integrated safeguards can be taken to mean meaningful detection capabilities across all plausible APs. Adherence to path level performance criteria for both timeliness and verification purposes can, in principle, ensure this.

By seeking paths where criteria are greatly exceeded, efficiencies can be identified. If standards were set at the level of individual proliferation actions, it is not clear how any such efficiencies could be recognized. Moreover, without recognizing the role of individual proliferation actions in acquisition paths, it is not clear how consistent and effective criteria could be crafted at the level of individual actions. Although any criteria could serve non-discriminatory objectives, it appears that path-level considerations can, in principle, avoid arbitrary standards and may be necessary to meet this key objective.

By focusing on path-level performance the Agency can explicitly take into account measures designed to detect undeclared activities including information analysis. Their contribution can be factored into the determination of path detection capability and thereby of criteria satisfaction.

Setting criteria for classes of paths can simplify the use of such criteria (see Example 3 above). Rather than having to determine criteria for individual paths, paths sharing common characteristics can be grouped together and assigned the same performance criteria. A few groups, and thereby a few sets of criterion, can cover all paths.

An analogy exists in current safeguards. Current safeguards implementation parameters are set based on the sensitivity of the material. For example, timeliness goals and thereby inspection frequencies are set depending on whether material is unirradiated direct-use material (one month), irradiated direct-use material (three months) or indirect-use material (one year). The more sensitive the material class, the more stringent the goal. The same principle could apply in determining detection parameters for classes of paths.

Measuring Safeguards Performance

When performance criteria are set, an evaluation must be done to ensure that the criteria will be met by a proposed safeguards approach. Care must be taken to demonstrate that, taken as a whole, the approach will meet the safeguards standards specified in the criteria. Moreover, this should be done before the approach is adopted. The APs to be detected must be complete and consistent in order to provide a common basis for comparing alternate approaches.

The Integrated Safeguards Evaluation Methodology (ISEM) was developed with these issues in mind.⁶ The ISEM provides internal consistency in the evaluation by defining a common set of acquisition paths and concealment possibilities. In ISEM, all APs relative to the scope of a given integrated safeguards proposal are selected for evaluation. This allows for the evaluation of limited proposals such as a single facility or for proposals that address an entire State.⁷ For a State-wide proposal, all paths would be selected.

The detection capabilities for individual proliferation actions—termed acquisition path segments in ISEM—are then assessed for the proposal considering possible concealment opportunities. The independent capabilities for each AP segment are then aggregated to determine path-level detection capabilities.

Conceptually, a path-level result for detection capability can be viewed as a cumulative probability distribution for time of AP detection.⁸ Performance criteria in effect place constraints on the allowed distribution for detection time. An approach may fail in that it does not provide a sufficient

detection capability to meet timeliness objectives. Alternatively, the approach may not provide high enough detection probability over time so that a credible level of assurance is attained.

Several elements may contribute to the determination of this distribution. For example, the random sampling used during inspections introduces a probabilistic factor. With random announced and unannounced inspections envisioned under integrated safeguards, the timing of inspections also becomes probabilistic. Both factors impact the likelihood that a given proliferation action would be detected within a given time period. Whether constraints are applied at the path or the segment level, ISEM can be used to determine whether criteria are being met.

Using Performance Criteria in a State-wide Context

Performance criteria could, in principle, be used by the Agency in a variety of ways. The options for criteria use range from developing initial, generic integrated safeguards approaches for typical fuel cycles to the continual use of the criteria to guide safeguards implementation in State-specific cases.

In the case of State-specific applications, the flexibility of a performance-based approach may be most beneficial. The same basic approach may present different costs in different States or present different impacts to the Agency or the State in question. In some cases, a “standard” Agency approach, if it were developed, may not be feasible in a State. With defined performance criteria and a consistent means for evaluating performance, a variety of acceptable approaches could be developed by the Agency. From this menu, the option that is most acceptable to the State may be selected.

Conclusions

The formal adoption of performance criteria would represent a significant, structural change in international safeguards. However, given the principles underlying the development of the current prescriptive criteria, establishing and using performance criteria would not be an entirely unfamiliar process for the Agency.

What may be more disconcerting is the potential loss of the prescriptive criteria that have been developed over decades of safeguards implementation. Such criteria form the basis for States’ expectations of the international safeguards system. However, the old criteria are not always relevant to integrated safeguards, and, from the outset, it has been recognized that under integrated safeguards the actual measures used in individual States may differ.

At this time of great change, confidence in the international safeguards system could be provided by acceptance of a rigorous and transparent process for safeguards development rather than a prescribed set of actions. The effectiveness as well as the consistency of safeguards implementation across States are both important issues. A safeguards system that sets performance criteria and has a well understood and widely accepted means for demonstrating that the criteria are being met could be an acceptable alternative to new prescriptive criteria. If prescriptive, baseline approaches for integrated safeguards are ultimately still desired, they can be developed using the same State-wide, performance-based approach.

Performance criteria could facilitate the successful attainment of a number of objectives for integrated safeguards. Performance criteria would provide a unifying principle for safeguards

implementation applicable to all States. Combined with a credible safeguards evaluation methodology, whether ISEM or any other methodology, they could be used to demonstrate and then optimize the effectiveness and efficiency of a safeguards approach. Doing so would allow the Agency to capture the benefits of flexible implementation without sacrificing performance.

¹ “Integrated Safeguards: Thinking About Implementation Criteria,” presented at the INMM-ESARDA Workshop on Science and Modern Technology for Safeguards, Tokyo, Japan, November 13-16, 2000.

² “Strengthening the Effectiveness and Improving the Efficiency of the Safeguards System and Application of the Model Protocol,” International Atomic Energy Agency, Vienna, 16 August 2000, GC(44)/12, p. 8

³ “IAEA Safeguards: Implementation at Nuclear Fuel Cycle Facilities,” International Atomic Energy Agency, Vienna, 1985, IAEA/SG/INF/6, ISBN 92-0-179085-6, p. 3

⁴ The notion of “verifying” the absence of undeclared activities is problematic as it introduces the notion of proving a negative. The term verification might better be reserved for inspections involving declared materials only. For entirely clandestine acquisition paths, timely detection may be the only relevant safeguards objective.

⁵ Compiled from GC(44)/12 (referenced above). The definition of integrated safeguards can be found on page 3 of this document. The basic principles of integrated safeguards are listed on page 8.

⁶ A more detailed description of the methodology is provided in: “Integrated Safeguards Evaluation Methodology (ISEM): Phase 2/Rev. 1: Concept Development,” Joseph F. Pilat, Kory W. Budlong Sylvester, George W. Eccleston, William D. Stanbro, Kenneth E. Thomas, Jim Larrimore, and Myron Kratzer, Los Alamos National Laboratory document LA-UR-00-1366, (March 2000). See also, “Illustrative Application of the Integrated Safeguards Evaluation Methodology: ISP-1/Rev. 1 and Current Safeguards,” Kory Budlong-Sylvester, Joseph F. Pilat, George W. Eccleston, William D. Stanbro, and Kenneth E. Thomas, Los Alamos National Laboratory document LA-UR-00-1362, March 2000.

⁷ For a discussion of the possible uses of ISEM, see Joseph F. Pilat and Kory Budlong Sylvester, “Possible Roles for the Integrated Safeguards Evaluation Methodology (ISEM),” presented at the 42nd Annual Meeting of the Institute of Nuclear Materials Management, Indian Wells, California, July 15-19, 2001; published in this proceedings.

⁸ A cumulative distribution $F(x)$ describes the probability that a random variable will take on a value less than or equal to x .